

Appendix A—Survey Methodology

This report contains data on state library agencies in the 50 states and the District of Columbia for state fiscal year (FY) 2000. The data were collected through the State Library Agencies (StLA) Survey, conducted by the National Center for Education Statistics (NCES). The reporting unit for the survey is the state library agency.

Survey Instrument

The StLA data are collected over the Internet via a Web-based reporting system. The Web survey application includes a user guide explaining the survey application, the data entry form and instructions, a 'Help' function which links data items to their definitions, and an edit check tool. The Web survey was designed to minimize response burden, to improve the timeliness and quality of the data, and to require minimal or no edit follow-up for data problems.

Survey Release Date, Editing, and Follow-up

The StLA Survey was released on the Web to state library agencies for data entry on October 16, 2000. The survey had a due date of February 16, 2001. The U.S. Census Bureau was the data processing agent for the survey through an interagency agreement with NCES. Census administered the Web survey system, provided technical support, conducted nonresponse follow-up, and produced the data files and tabulations of the data. Census edited the state data submissions, with additional review by NCES.

To reduce response burden, the survey was transmitted with pre-entered prior-year data for items where the data are not expected to change annually—Most of Parts A through E and N, or about half of all survey items. The respondent was requested to review the pre-entered data and update any information that had changed from the previous year. All other data cells were left blank for the respondent to fill in, not update. The respondent was instructed to answer -1 to any numeric item if they could not provide the data. The Web application required a response to all alpha-numeric items, except items that could legitimately be left blank due to skip patterns. The respondent could not "lock" his/her data file to signify completion of the survey unless these conditions were met. A zero (0) is a reported response and indicates the state library agency had none of the item. Missing data were imputed. See the section below on "Imputation" for a discussion of the imputation methodology.

An edit check tool, built into the Web application, alerted the respondent to questionable data during the data entry process through interactive "edit check warnings" and through edit check reports which could be viewed on-screen or printed. The edit check program enabled the respondent to submit an edited data file to NCES which usually required little or no follow-up for data problems. The edit check tool includes four types of edits:

1. Relational edit checks. This is a data consistency check between related data elements. For example, an edit check message is generated if the state library agency is designated as a federal depository library but does not indicate the type of federal depository library.
2. Out-of-range edit checks. This is a comparison of data reported for an item to the "acceptable range" of numeric values. For example, an edit check message is generated if annual circulation transactions per annual library visits is less than 0.5 circulation transactions per visit.
3. Arithmetic edit checks. This is an arithmetic check comparing a reported total to the sum of its parts. For example, an edit check message is generated if total operating expenditures is not equal to the sum of its parts (total staff expenditures, collection expenditures, and other operating expenditures).
4. Blank/zero/invalid edit checks. This is a check of reported data against acceptable values. For example, an edit check message is generated if book/serial volumes is 0 or blank.

The preliminary data file and draft tables were reviewed by the State Library Agencies Survey Steering Committee, NCES, and Census for data quality. Based on this review, states with questionable data were contacted to request verification or correction of the data.

Imputation

Missing data were imputed using one of 4 methods: the zero rule, the growth rule, regression modeling, or the sum rule. These methods were used in order—the zero rule first, then the growth rule, then regression modeling, and finally the sum rule. These methods are explained below.

- ? ? **The zero rule.** If the state did not report a value for FY 2000 and the value is 0 for FY 99, then the value for FY 2000 was set to 0. This rule was applied first, on the assumption that there was no change from the prior year. The 0 in the prior year could be an imputed value.
- ? ? **The Growth Rule.** If the state did not report a value for FY 2000 and the value for FY 99 was greater than 0, the growth rate from FY 99 to FY 2000 was calculated for all states that reported data greater than 0 in both years. The median of the growth rates was then calculated and applied to the state's FY 99 data to obtain an estimate for FY 2000. The growth rate could be applied to an FY 99 imputed value. Note: The growth rule looked at values for FY 99 only, as states that cannot report a particular item tend to have ongoing problems reporting the item, so earlier data were not considered to be useful.
- ? ? **Regression Modeling.** If the state did not report a value for FY 2000 and there was no comparable value for FY 99 (e.g., a new item was added to the survey), regression was used. The regression model used only the FY 2000 data file. Regression was used to impute adult literacy and family literacy (ADFAMLIT),¹ a revised item on the FY 2000 file that did not directly correspond to the FY 99 item (adult literacy). Total literacy program support staff (LITPRSED), total state expenditures for financial assistance to libraries (AIDPLSB), and total LSTA expenditures for services to children in poverty (SERPOVXU) were chosen as the predictor items reported by all states. Due to response scaling problems, a log transform was used on ADFAMLIT. The model was fitted to states that reported a positive value for ADFAMLIT. The missing value of ADFAMLIT was then replaced with the regression-predicted value.
- ? ? **The Sum Rule.** When the details of a total and the total were missing, the details were imputed by the zero rule, the growth rule or regression modeling. The total was then imputed by adding up the details.

Survey Data Items

The FY 2000 survey collected data on 423 items, including state library agency identification, governance, public service hours, service outlets, collections, library service transactions, library development transactions, services to other libraries in the state, allied operations, staff, income, expenditures, and electronic services and information. The data items and definitions are provided in the survey instrument in appendix D.

Universe

The state library agencies in the 50 states and the District of Columbia (51 total) comprise the survey universe.

¹ The words in parentheses are the variable names on the data file.

Response Rate

Unit Response. The FY 2000 StLA Survey achieved a 100 percent unit response rate.

Item Response. Most items had a 100 percent response rate. Items with less than a 100 percent response rate are listed below.

| Item number | Name | Response rate |
|----------------|---|------------------|
| 114 | Library visits | 90.2 |
| 223a | Statewide database licensing expenditures (federal) | 98.2 |
| 223b | Statewide database licensing expenditures (state) | 98.2 |
| 223c | Statewide database licensing expenditures (other) | 98.2 |
| 223d | Statewide database licensing expenditures (total) | 98.2 |
| 239 | Pre-kindergarten learning ("readiness for school") | 98.2 |
| 240 | Adult literacy and family literacy | 98.2 |

Reporting Period. The FY 2000 StLA Survey requested data for state fiscal year 2000, except for Part B—Governance and Part J—Staff which requested data as of October 1, 2000. The fiscal year of most states was July 1, 1999 to June 30, 2000. Exceptions were New York (April 1, 1999 to March 31, 2000); Texas (September 1, 1999 to August 31, 2000); and Alabama, the District of Columbia, and Michigan (October 1, 1999 to September 30, 2000).

Using the Data to Make Comparisons. Missing data were imputed. The imputations are included in national totals but suppressed at the state level in the tables in this report. Imputations were first included on the FY 99 survey file; users should therefore take into consideration that the data are not strictly comparable to data prior to FY 99, which rely only on reported data.

State comparisons should be made with caution because states vary in their fiscal year reporting periods, as indicated above, and may vary in their interpretation of the survey definitions. The District of Columbia, while not a state, is included in the survey. Caution should be used in comparing District of Columbia data with state data.

(page intentionally blank)

